

# Screening of Structures in Highly Fluorinated Anatase $\text{TiO}_2$ by Computer Simulations

Dario Corradini,<sup>1,2</sup> Damien Dambournet,<sup>1,2</sup> and Mathieu Salanne<sup>1,2,3,\*</sup>

<sup>1</sup>*Sorbonne Universités, UPMC Univ Paris 06, UMR 8234, PHENIX, Paris, France*

<sup>2</sup>*CNRS, UMR 8234, PHENIX, Paris, France*

<sup>3</sup>*RS2E, FR CNRS 3459, France*

(Dated: March 24, 2014)

A highly fluorinated anatase lattice has been recently reported, providing a new class of materials whose general chemical formula is  $\text{Ti}_{1-x}\square_x\text{X}_{4x}\text{O}_{2-4x}$  ( $\text{X}^- = \text{F}^-$  or  $\text{OH}^-$ ). Both experiments and simulations have revealed the presence in the material of cationic vacancies ( $\square$ ) and of three different environments for the fluorine,  $\text{F} - \text{Ti}_1$ ,  $\text{F} - \text{Ti}_2$  and  $\text{F} - \text{Ti}_3$ . In order to better characterize the structural features of the material and the effect of the different F environments, we here apply a computational screening procedure. After deriving a polarizable force-field from DFT simulations, we screen in a step-wise fashion a large number of possible configurations differing in the positioning of the titanium vacancies and of the fluorine atoms. At every step only the 10 % configurations with the lowest energy are retained. *NPT* Molecular Dynamics simulations at ambient conditions are performed on the configurations selected by the screening procedure. The method allows to predict the most favorable structural arrangements of the material and at the same time to discern structural characteristics impossible to disentangle by experimental techniques only.

Titanium dioxide is a widely studied material due to its promising applications in many fields among which photocatalysis, green chemistry and energy storage [1–8]. Interest in the polymorphs of  $\text{TiO}_2$ , among which the naturally occurring rutile, anatase and brookite, was as well sparked by the possible application as anodic material in Li ion batteries [9–14]. Fluorinated  $\text{TiO}_2$  has also been investigated [15–21] since the presence of F in the compound might improve the sought characteristics of the material [17] or stabilize the highly reactive {001} facets of the anatase crystal [20, 22]. The nature of the fluorinated compound depends strongly on the fluorination technique employed [18, 23]. So far the stabilization of fluorine within the anatase lattice of  $\text{TiO}_2$  has been poorly understood, owing to the structural complexity of the fluorinated compound.

Recently, a novel synthesis technique conducted in our laboratory [24] led to the preparation of a highly fluorinated anatase material in which fluorine replaces oxygen atoms in their lattice sites and the resulting charge deficiency is compensated by the formation of a cationic vacancy every four substitutions. The material obtained has thus the general formula  $\text{Ti}_{1-x}\square_x\text{X}_{4x}\text{O}_{2-4x}$ , where the symbol  $\square$  is used to represent a cationic vacancy and X is either a  $\text{F}^-$  ion or an  $\text{OH}^-$  group (the amount of  $\text{F}^-$  may vary depending on the synthesis condition). Elemental analysis and synchrotron diffraction revealed the existence of more than 20% cation vacancies, in fact the stoichiometric formula  $\text{Ti}_{0.78}\square_{0.22}\text{X}_{0.88}\text{O}_{1.12}$  was assigned to the material. By using  $^{19}\text{F}$  NMR spectroscopy it was also possible to discern three different coordination modes for the F atoms in the material:  $\text{F} - \text{Ti}_1$ ,  $\text{F} - \text{Ti}_2$  and  $\text{F} - \text{Ti}_3$ , highlighting the complex structural arrangement present in the material.

In this Letter, we report the results of a computational study of the fully-fluorinated material (e.g.  $\text{Ti}_{0.78}\square_{0.22}\text{F}_{0.88}\text{O}_{1.12}$ ) performed in order to better characterize its structural features and the effects induced by the three different environments around the F atoms. The enormous number of possible structural arrangements of the vacancies and of the F atoms in the anatase structure render the problem untreatable directly by ab initio simulations. Therefore we apply a screening procedure on the possible configurations in the spirit of the emerging high-throughput techniques, see e.g. Refs. [25–27], by using classical Molecular Dynamics (MD). Several force-fields were previously proposed for pure  $\text{TiO}_2$  [28–34]. Here we derive a new *polarizable* force-field valid for the pure phase as well as for the fluorinated material by extracting the force-field parameter from DFT simulations, via a force and dipole fitting procedure well-established in our group (see for example Refs. 35 and 36). The details of the force-field are discussed in the Appendix. The parameters of the force-field obtained from DFT simulations are reported in Appendix Table I and II. The classical polarizable force-field includes a damped Born-Mayer-Huggins-Fumi-Tosi term, a Coulombic term and a polarization term described by making use of the induced dipoles.

In order to generate and assess fluorinated samples starting from the pure  $\text{TiO}_2$  anatase, we apply a screening procedure, similar for example to what done by Wilmer *et al.* for metal-organic frameworks (MOFs) [37]. At the fixed target composition  $\text{Ti}_{0.78}\square_{0.22}\text{F}_{0.88}\text{O}_{1.12}$  we consider samples containing: a) only F in the environment  $\text{F} - \text{Ti}_2\square_1$ ; b) F in both environments  $\text{F} - \text{Ti}_2\square_1$  and  $\text{F} - \text{Ti}_3$ ; c) F in all possible environments  $\text{F} - \text{Ti}_1\square_2$ ,  $\text{F} - \text{Ti}_2\square_1$ , and  $\text{F} - \text{Ti}_3$ . We leave the ratio of F in the different environments free to vary at random (albeit with constraints, see below). The starting fluorinated structures are generated from the  $4 \times 4 \times 2$  pure anatase  $\text{TiO}_2$  structure [38] ( $\text{Ti}_{128}\text{O}_{256}$ ) leading to a system thus

\* Corresponding Author. E-mail: mathieu.salanne@upmc.fr

composed:  $\text{Ti}_{100}\square_{28}\text{F}_{112}\text{O}_{144}$ . These configurations are generated applying the following criteria:

- a) for the case with only F in the environment  $\text{F} - \text{Ti}_2\square_1$ , we erase 28 Ti ions at random with the constraint that they must not be adjacent to a previously selected Ti either at the same height along [001] or on different heights. At each cationic vacancy thus created we assign a random number of localized F  $N_F^i$  between 1 and 6 and then for each vacancy we replace the closest  $N_F^i$  O with F with the constraint that the total number of F must correspond to the stoichiometric number,  $N_F = 112$ . All F are therefore either found at the shortest  $\square - \text{F}$  distance  $\simeq 1.93$  Å (at “equatorial” positions in the original  $\text{TiO}_6$  octahedron), or at second shortest  $\square - \text{F}$  distance  $\simeq 1.98$  Å (at “axial” positions in the original  $\text{TiO}_6$  octahedron).
- b) for the case with F in both environments  $\text{F} - \text{Ti}_2\square_1$  and  $\text{F} - \text{Ti}_3$ , we create the vacancies with same procedure described in a) and then randomly substitute 112 O with 112 F. The constraint of having non-adjacent vacancies is sufficient to ensure there are no F with environment  $\text{F} - \text{Ti}_1\square_2$  in the system.
- c) for the case with in all possible environments  $\text{F} - \text{Ti}_1\square_2$ ,  $\text{F} - \text{Ti}_2\square_1$  and  $\text{F} - \text{Ti}_3$  we erase 28 Ti ions at random with no constraints on the creation of adjacent vacancies and we randomly substitute 112 O with 112 F. We impose that all F and O must be attached to at least one Ti. The configurations not respecting this criterion are discarded.

For the three cases a), b) and c) a screening procedure is then initiated. The protocol of the screening procure is as follows:

- 1) we perform single-point energy calculations on  $\sim 1.5 \cdot 10^5$  configurations generated for each case a); b) and c) as described above; we sort the configurations according to their ascending energy and we retain the  $\sim 1.5 \cdot 10^4$  configurations with the lowest energy for the following step.
- 2) we perform 0 K geometry optimizations of the atomic positions, keeping the length of the cell vectors fixed; we sort the configuration according to their ascending energy and we retain at maximum the  $\sim 1.5 \cdot 10^3$  configurations with the lowest energy for the following step. The configurations crashed during the run are discarded.
- 3) we perform 0 K cell optimizations allowing the contemporary adjustment of the atomic positions and of the lengths of the cell vectors, while keeping the box angles fixed at  $\alpha = \beta = \gamma = 90^\circ$ ; we sort the configuration according to their ascending energy and we retain at maximum the  $\sim 1.5 \cdot 10^2$  configurations with the lowest energy for the following

step. The configurations crashed during the run are discarded.

- 4) we temper the configurations performing 10 ps  $NVT$  runs at finite temperatures from  $T_1 = 25$  K to  $T_{12} = 300$  K, every  $\Delta T = 25$  K. We control the temperature through the Nosé-Hoover thermostat [39, 40] with the time constant  $\tau_T = 1$  ps. The 15 configurations with the lowest energy at  $T_{12} = 300$  K are retained for the following step.
- 5.1) we equilibrate the remaining samples for 100 ps in the  $NVT$  ensemble at  $T = 300$  K, with  $\tau_T = 1$  ps.
- 5.2) we equilibrate the remaining samples for 200 ps in the  $NPT$  ensemble at  $T = 300$  K ( $\tau_T = 2$  ps) and at ambient pressure. Pressure is controlled by using the Martyna barostat [41] with  $\tau_P = 4$  ps.
- 5.3) for the remaining samples, we perform a 1 ns production run in the  $NPT$  ensemble. The time constants of the thermostat and the barostat are set respectively to  $\tau_T = 10$  ps and  $\tau_P = 10$  ps.

At the end of the screening procedure we are left with 15 configurations from group a) i.e. initially containing only  $\text{F} - \text{Ti}_2\square_1$ , 8 configurations from group b) i.e. initially containing only  $\text{F} - \text{Ti}_2\square_1$  and  $\text{F} - \text{Ti}_3$  and 10 configurations from group c), i.e. initially containing F in all the three environments.

In Fig. 1 we show the the relative frequency histograms built considering all the potential energies obtained at the end of each screening step from 1) to 4) for all the cases a); b) and c). Together we show the sorted energies of the configurations retained to be run in the following step. We stress here that we consider the ratios of  $\text{F} - \text{Ti}_1$ ,  $\text{F} - \text{Ti}_2$  and  $\text{F} - \text{Ti}_3$  as the ones calculated for the starting unrelaxed configuration. We will see in the following that their relative amount can change during the simulations. Looking at Fig. 1 we can immediately observe that the configurations prepared with only  $\text{F} - \text{Ti}_2\square_1$  tend always to be the lowest in energy at all steps. This is consistent with the stabilization of the vacancies induced by the presence of the F atoms around them as observed in Ref. [24]. We can also observe that upon relaxation of the structures and upon tempering of the configurations the difference in potential energy between the best configurations initially containing only  $\text{F} - \text{Ti}_2\square_1$  and the best ones prepared with F in all possible environments diminishes dramatically, see Fig 1(b,d,f,h). The unrelaxed structures at 0 K show a difference in energy from  $\simeq 1.2$  to  $\simeq 2.1$  eV/f.u. (where we consider 128 formula units, f.u.). After the cell relaxation this  $\Delta E$  reduces to  $\simeq 45.5$ – $81.7$  meV/f.u. Finally the tempered structures, at  $T = 300$  K, show a  $\Delta E \simeq 22.4$ – $38.2$  meV/f.u. This difference is then comparable to the thermal energy  $k_B T$ , which equals 25.9 meV at  $T = 300$  K.

Therefore we see that the configurations initially prepared with all F environments are stabilized both by allowing the relaxation of the structure and by the effect

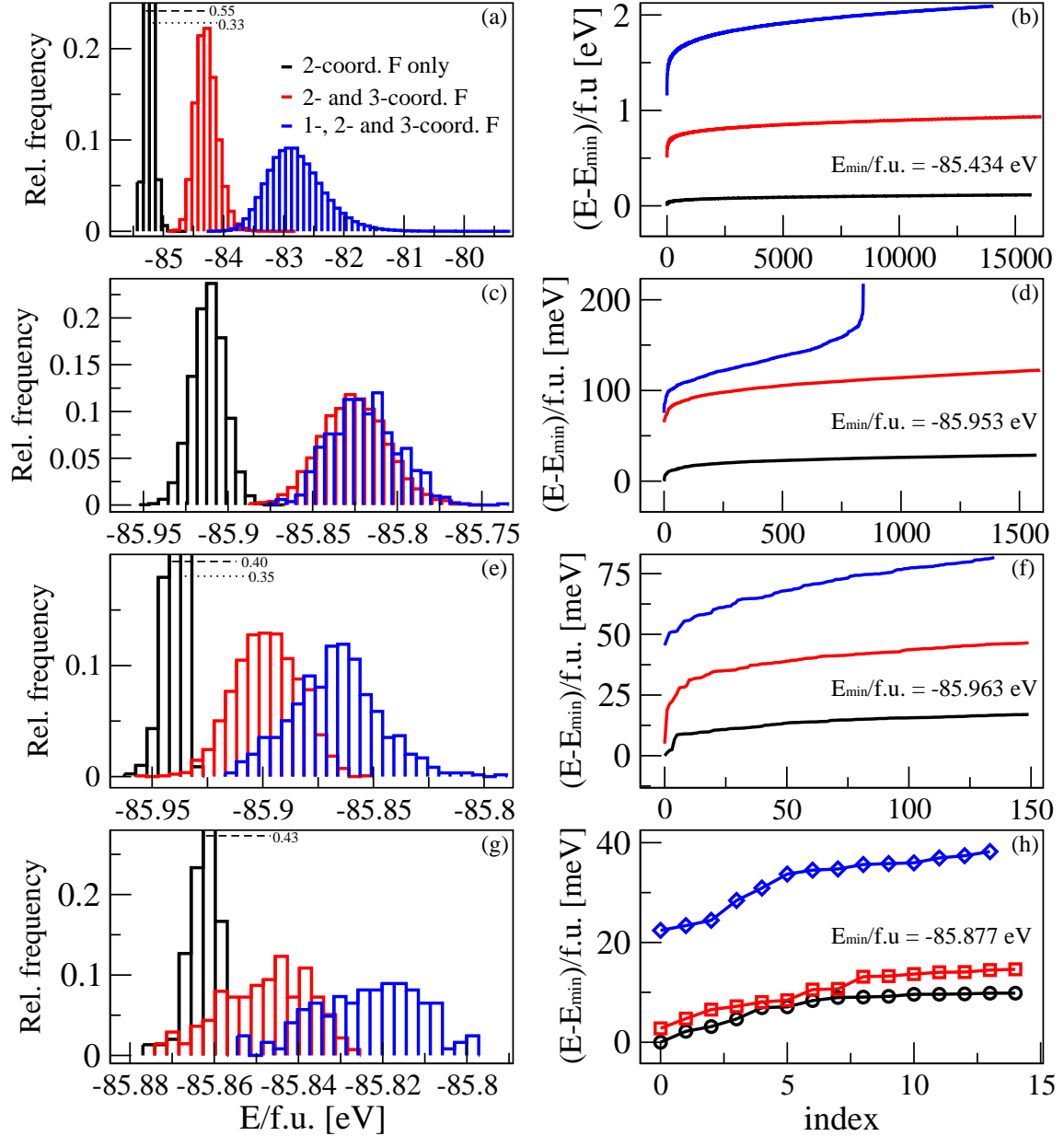


FIG. 1. The panels on the left show the relative frequency histograms of the energies calculated for all the configurations tested at each step, while the panels on the right show the sorted energies of the configurations retained after each step. The energies shown are: (a,b) at 0 K for the starting unrelaxed structures; (c,d) at 0 K after the optimization of the atomic positions; (e,f) at 0 K after the optimization of the atomic positions and cell vector lengths; (g,h) at 300 K after the tempering from 25 to 300 K. All steps are performed for configurations initially containing F – Ti<sub>2</sub>□<sub>1</sub> only (black), F – Ti<sub>2</sub>□<sub>1</sub> and F – Ti<sub>3</sub> (red) and F – Ti<sub>1</sub>□<sub>2</sub>, F – Ti<sub>2</sub>□<sub>1</sub> and F – Ti<sub>3</sub> (blue).

of temperature. We underline here the importance of the thermal treatment since the lowest energy difference at 0 K after the cell optimization remains larger than 45 meV/f.u. The case initially containing only, F – Ti<sub>2</sub>□<sub>1</sub> and F – Ti<sub>3</sub> is intermediate between the other two and the difference in energy between these configurations and the ones containing initially only F – Ti<sub>2</sub>□<sub>1</sub> becomes very small at  $T = 300$  K.

Capturing these thermal effects in a generic, entirely *ab*

*initio* based high-throughput procedure would be impossible. Generally, such studies involve static calculations only since performing molecular dynamics simulations is computationally too expensive (and, very often, useless). Nevertheless, it is interesting to test whether our selected configurations (i.e. those remaining at the end of all the screening procedure) would also have been selected if *ab initio* static calculations had been performed. To test this, we take their initial structures and calculate their

DFT energy. Then we take the same number of random configurations from the pools a); b); c). We find that (i) the configurations given by the classical screening all have a lower DFT energy than the ones taken at random and (ii) the difference in energy between configurations containing only  $F - Ti_2\Box_1$  and the ones containing  $F - Ti_2\Box_1$  and  $F - Ti_3$  are reproduced quantitatively in the DFT calculations while the energy difference with the configuration containing F in all environment is slightly lower in DFT. The results of this validation are shown in Appendix Fig. 5.

Next, we analyze how the initial structural arrangements correlate with the energy of the configurations. The results are shown in Fig. 2. For the configurations containing only  $F - Ti_2\Box_1$ , see Fig. 2(a), we see that lower energies correlate with a higher fraction of F in the “equatorial” configuration. In fact in the anatase structure each Ti (or vacancy) has 4 “equatorial” neighboring O (or F) at a shorter Ti (or  $\Box$ )-O (or F) distance and 2 “axial” neighboring O (or F) at a  $\sim 2.5\%$  longer Ti (or  $\Box$ )-O (or F) distance. As already indicated by the DFT calculations shown in Ref. 24, having the F closest to the vacancy stabilizes the structures. We note that the *a priori* probabilities of having equatorial F is 86% and the one of having axial F is 14% with the criteria mentioned above for building the pool a). In the same graph we also report the initial energies of the best configurations given at the end by the screening procedure. We observe that they are found in the lowest energy–higher fraction of equatorial F corner.

For the case in which we have  $F - Ti_2\Box_1$  and  $F - Ti_3$ , see Fig. 2(b), lower energies correlate with a higher fraction of  $F - Ti_2\Box_1$ , again consistently with the observed stabilization of the vacancies by the F. In this case the *a priori* probabilities of having  $F - Ti_2\Box_1/F - Ti_3$  is 65.6/34.4 % according to the criteria used to build the pool b). In this case the final screening configurations are found closer to the average F speciation values rather than at the highest  $F - Ti_2\Box_1$  relative compositions.

Also for the case containing all the three F environments, see Fig. 2(c), lower energies correlates with the percentage of  $F - Ti_2\Box_1$ . Here the *a priori* probabilities of having  $F - Ti_1\Box_2/F - Ti_2\Box_1/F - Ti_3$  is  $\simeq 11/43/46$  %. The configurations surviving the screening procedure in this case are even more dispersed in terms of their  $F - Ti_2\Box_1$  content.

As a matter of fact the  $^{19}F$  NMR analysis on the experimental sample reveals at  $T = 150$  K relative ratios of  $F - Ti_1/F - Ti_2/F - Ti_3$  of 13/70/17 % [24], indicating therefore a very special F speciation. This is not inconsistent with our findings. In fact upon relaxing the structures and increasing their temperature we also find a large increase of the number of  $F - Ti_2$ . The F ratios after the dynamic runs is difficult to calculate exactly because the presence of disorder induces a large dependence of the results on the choice of the cutoff distance used for assigning whether a fluorine is coordinated to a titanium atom. Nevertheless we systematically observe

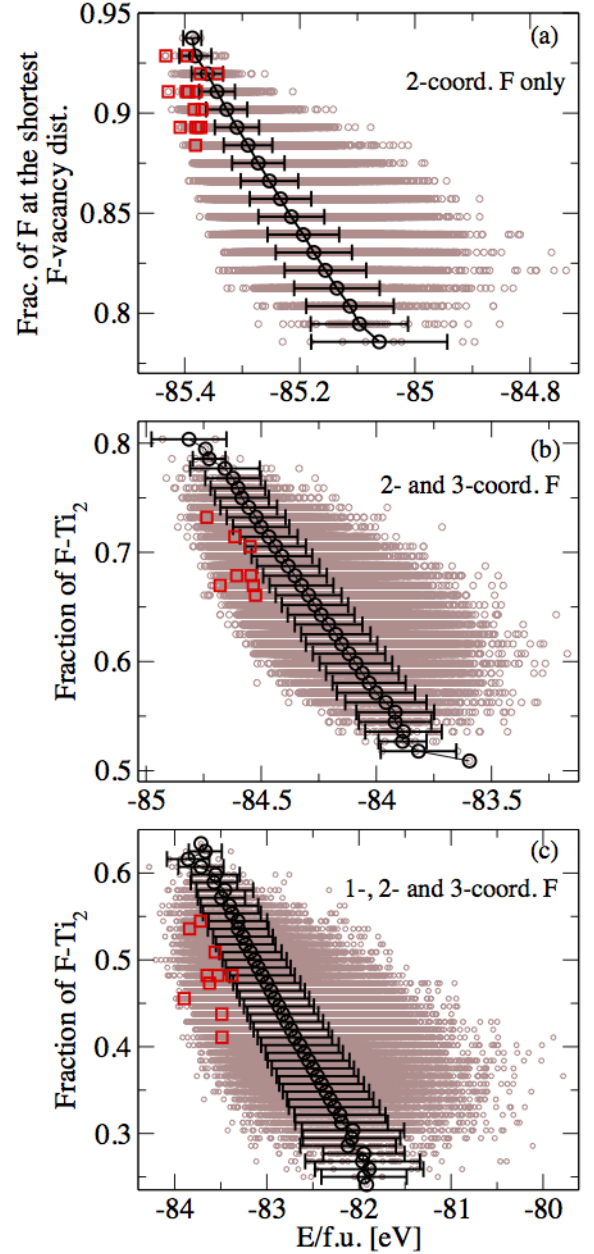


FIG. 2. For the initial configurations at 0 K (a) fraction of “equatorial” F atoms (at the shortest  $F - \Box$  distance) vs. energy of the configuration for the case containing  $F - Ti_2\Box_1$  only; (b) fraction of  $F - Ti_2\Box_1$  vs. energy of the configuration for the case containing  $F - Ti_2\Box_1$  and  $F - Ti_3$ ; (c) fraction of  $F - Ti_2\Box_1$  vs. energy of the configuration for the case containing  $F - Ti_1\Box_2$ ,  $F - Ti_2\Box_1$  and  $F - Ti_3$ . The points are represented as small brown circles. For each different value of the fraction of “equatorial” F (a) or  $F - Ti_2\Box_1$  (b, c) we calculate the mean (black circles) and the standard deviation (black bars) of the corresponding energies. We also report the values assumed at this stage by the configurations run at the final screening step (red squares).

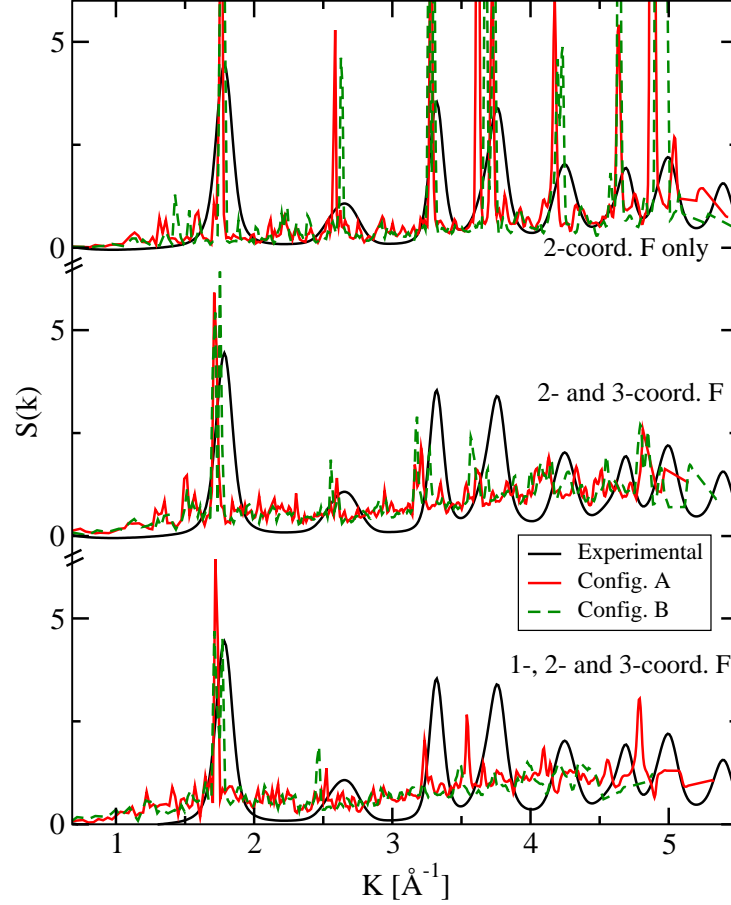


FIG. 3. Comparison of the structure factor  $S(k)$  at ambient conditions measured in experiments (black solid line) and calculated as described in the text from  $NPT$  runs of two configurations given by the screening procedure that we label A (red solid line) and B (green dashed line) for each case: initial  $F - Ti_2\Box_1$  only,  $F - Ti_2\Box_1$  and  $F - Ti_3$  and  $F - Ti_1\Box_2$ ,  $F - Ti_2\Box_1$  and  $F - Ti_3$ .

an increase in the number of 2-coordinated F after relaxing and heating. In the following we will consider a quite large F–Ti distance,  $d_{F-Ti} = 3.0 \text{ \AA}$  for evaluating the F– $Ti_1$ /F– $Ti_2$ /F– $Ti_3$  speciation for the final configurations (at ambient conditions). This is also justified by the fact that the Ti–F radial distribution function does not decay to 0 between the first and the second peaks (in the cases not containing only  $F - Ti_2\Box_1$  initially, see Fig. 4 in the following).

In order to compare the structural properties of the material as found in the experiments with our simulations, we compare the experimental x-ray structure factor  $S(k)$  to the one we calculate from our trajectories using the Ashcroft–Langreth partial structure factors according to the formula:

$$S_{\text{tot}}(k) = \frac{\sum_{\alpha\beta} \sqrt{x_\alpha x_\beta} f_\alpha(k) f_\beta(k) S_{\alpha\beta}(k)}{\sum_\alpha x_\alpha f_\alpha^2(k)} \quad (1)$$

where  $x_\alpha$  are the relative atomic concentrations and  $f_\alpha(k)$  are the  $k$ -dependent atomic scattering. They are calculated using the analytic approximation:

$$f_\alpha(k) = c_\alpha + \sum_{i=1}^4 a_{\alpha,i} \exp \left[ -b_{\alpha,i} \left( \frac{k}{4\pi} \right)^2 \right] \quad (2)$$

where the coefficients  $a_{\alpha,i}$ ,  $b_{\alpha,i}$  and  $c_\alpha$  are taken from Ref. [42] for  $O^{2-}$  and from Ref. [43] for  $Ti^{4+}$  and  $F^-$ . For each pool of configurations, we select the two trajectories that best reproduce the experimental  $S(k)$  and we label them A and B. The results are shown in Fig. 3. It is worth noting that in experiments part of the fluorine are replaced by hydroxide groups, which should of course affect the comparison. Nevertheless the position of the peaks and their intensities remain good indicators of the level of accuracy of our structure prediction methodology.

With the choice for the F–Ti distance mentioned above, we note the following for the speciation of the selected configurations. The F –  $Ti_1$ /F –  $Ti_2$ /F –  $Ti_3$  speciation of the initial (final) configurations is: initial F –  $Ti_2$  only configuration A 0.0/100.0/0.0 % (0.0/96.4/3.6 %), configuration B 0.0/100.0/0.0 % (0.0/100.0/0.0 %); initial F –  $Ti_2$  and F –  $Ti_3$  configuration A 0.0/71.4/28.6

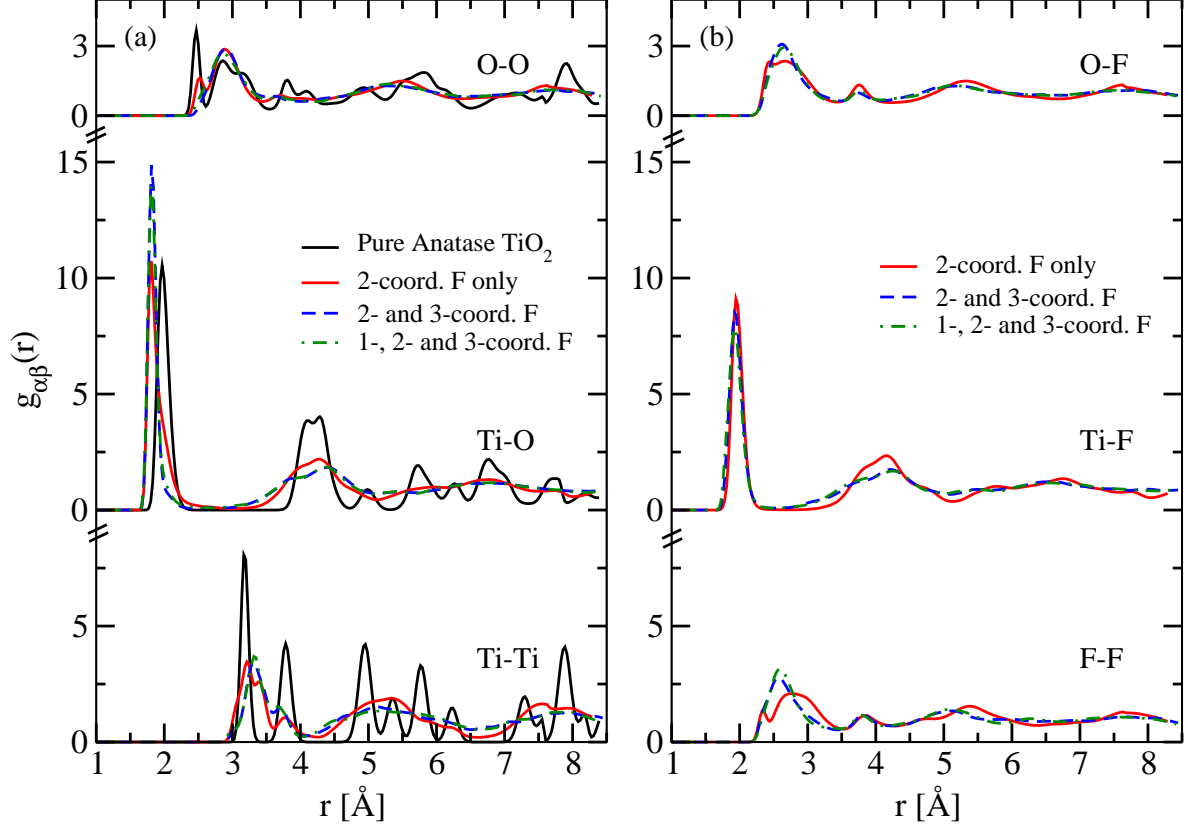


FIG. 4. Partial radial distribution functions  $g_{\alpha\beta}(r)$  at ambient conditions calculated from *NPT* simulations of pure  $\text{TiO}_2$  anatase (black solid line) and from simulations of one of the configurations given by the screening procedure (configuration labelled A of Fig. 3) for the cases: initial F –  $\text{Ti}_2\Box_1$  only (red solid line), F –  $\text{Ti}_2\Box_1$  and F –  $\text{Ti}_3$  (blue dashed line) and F –  $\text{Ti}_1\Box_2$ , F –  $\text{Ti}_2\Box_1$  and F –  $\text{Ti}_3$  (green dot-dashed line).

% (0.9/87.5/11.6 %), configuration B 0.0/73.2/26.8 % (2.7/85.7/11.6 %); initial F –  $\text{Ti}_1$ , F –  $\text{Ti}_2$  and F –  $\text{Ti}_3$  configuration A 12.5/47.3/40.2 % (4.5/90.2/5.3 %), configuration B 11.6/54.5/33.9 % (2.7/85.7/11.6 %). Thus we see that the classical potential used might have the drawback to exaggerate the percentage of 2-coordinated F atoms with respect to the experiment where the F speciation is found to be 13/70/17 % (albeit this was found at  $T = 150$  K).

Nonetheless we can observe in Fig. 3 that the agreement between the experimental and the calculated structure factor is quite good. The peak slight misalignment is due to an enlargement of the simulation box with the respect to the experiment. In fact one average we found a 2.6% increase of the lattice parameters for pure anatase, to be compared with the 0.5% increase in DFT calculations. It is clear that the intensity of the peaks for the configurations initially prepared with only F –  $\text{Ti}_2\Box$  is overestimated but in the other two cases the intensity of the peaks is almost in quantitative agreement with the experiment. The higher level of noise in the simulations

results is due to the relatively small size of the box compared to the experiment. The results described for the  $S(k)$  validate the study of the structure by our model.

Consequently, in Fig. 4 we show the partial radial distribution functions at ambient conditions  $g_{\alpha\beta}(r)$  for the simulated pure anatase  $\text{TiO}_2$  and for the configurations labelled A in Fig. 3. The effect of the disorder introduced by the vacancies is immediately evident looking at the  $g_{\text{Ti-Ti}}(r)$  and at the  $g_{\text{O-O}}(r)$ . The  $g_{\text{Ti-O}}(r)$  structure seems to be conserved to a large extent at least for the first two shells, although the effect of disorder is evident when looking for example at the region in between the first two peaks. When comparing the configuration initially having only F –  $\text{Ti}_2$  to the other two cases it is evident that the first conserves a greater degree of order with respect to the others. In fact all its  $g_{\alpha\beta}(r)$  present peaks and shells which are more defined. As we have seen in Fig. 3 however the experimental structure should be more similar to the one of the configurations prepared having initially also F –  $\text{Ti}_1$  or F –  $\text{Ti}_3$ . For those cases we can observe the similarity of the  $g_{\text{Ti-O}}(r)$  and



$g_{\text{Ti-F}}(r)$ , confirming the substitutional nature of the F on the O sites also at ambient temperature. However we also note that the Ti – O first neighbor distance become smaller in the fluorinated samples with the first peak of the  $g_{\text{Ti-O}}(r)$  centered at  $\simeq 1.81$  Å, to be compared with  $\simeq 1.94$  Å for  $g_{\text{Ti-F}}(r)$ . Connected to this, the first peak of the  $g_{\text{O-O}}(r)$  is at a larger distance ( $\simeq 2.9$  Å) with respect to the first peak of the  $g_{\text{F-F}}(r)$  ( $\simeq 2.6$  Å). We also observe that the  $g_{\text{O-O}}(r)$  is fairly similar to the  $g_{\text{F-F}}(r)$ . Summarizing, we find that fluorine is on average farther from Ti than O and closer to O than other oxygen ions, while the distance of fluorine from other F ions is found to be similar to the distance between F and O.

In conclusion, in order to characterize fluorinated anatase  $\text{Ti}_{0.78}\square_{0.22}\text{F}_{0.88}\text{O}_{1.12}$  as produced in experiments [24], we develop a classical polarizable force-field from DFT simulations. With the derived force-field we initiate a screening procedure on three large pools of configurations containing a) only F –  $\text{Ti}_2\square_1$ ; b) F –  $\text{Ti}_2\square_1$  and F –  $\text{Ti}_3$ ; c) F –  $\text{Ti}_1\square_2$ , F –  $\text{Ti}_2\square_1$  and F –  $\text{Ti}_3$ . We run *NPT* simulations at ambient conditions on the best configurations left after the screening procedure. The comparison of structure factor calculated from the our simulations to the one measured in experiments reveals a very good agreement of the structural features. We find that upon relaxation and heating the number of 2-coordinated F increases, consistently with experimental results. The study of the partial radial distribution functions shows that the average first neighbor distance is shorter for Ti-O than Ti-F, the reverse is true for O-O and O-F. Our force-field combined to the screening procedure allowed us (i) to select the best configurations starting from a very large pool of possible configurations; (ii) to reproduce the experimental structure to a good extent, (iii) to study details of the partial atomic structure unaccessible in experiments. Furthermore, this procedure appears promising in general for the study of materials presenting complex structural arrangements.

## ACKNOWLEDGEMENTS

We thank François-Xavier Coudert for introducing us to the screening techniques. We thanks Benjamin J. Morgan and Paul A. Madden for discussions. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° 321879.

## APPENDIX

Appendix Figure 5 shows the validation of the screening procedure performed by DFT calculations as described in the text. The derivation of the polarizable classical force field is also described in the Appendix. The

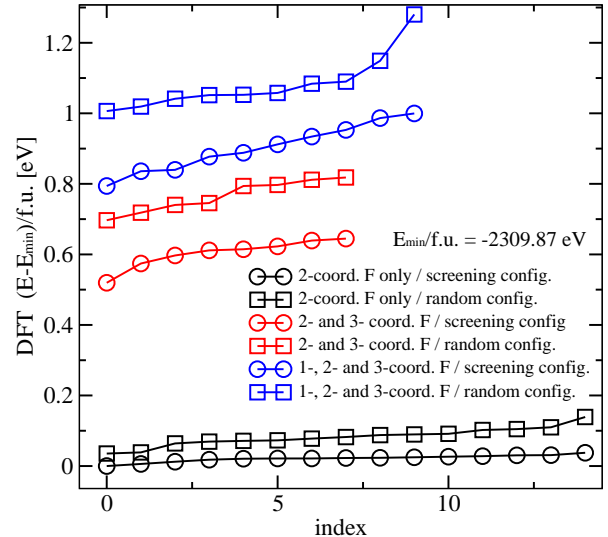


FIG. 5. For the configurations left at the end of the screening procedure, we take their initial unrelaxed structures and calculate their DFT energy (circles). Those are compared with the DFT energies of the same number of configurations taken at random (squares) from the initial pools of configuration with F –  $\text{Ti}_2\square_1$  only (black); F –  $\text{Ti}_2\square_1$  and F –  $\text{Ti}_3$  (red) and F –  $\text{Ti}_1\square_2$ , F –  $\text{Ti}_2\square_1$  and F –  $\text{Ti}_3$  (blue).

parameters of the potential are reported in Appendix Tables 1 and 2.

## Classical model and derivation

We describe the interaction potential between the ions by a classical polarizable force-field whose parameters we derive from *ab initio* DFT simulations. Details on how to extract the parameters of the classical potential from DFT simulations are reported in earlier works [35, 36]. Here in the following we give the details of the DFT simulations performed and we illustrate the details of the chosen model for the ions. We then present the parameters derived by fitting the forces and the dipoles obtained by the classical model to the ones obtained in DFT.

## DFT calculations

The initial structures are obtained from the pure anatase  $\text{TiO}_2$   $3 \times 3 \times 1$  structure (36  $\text{TiO}_2$  units, total 108 atoms) creating vacancies and substituting F to O according to the stoichiometric formula  $\text{Ti}_{1-x}\square_x\text{F}_{4x}\text{O}_{2-4x}$ . The DFT simulations are performed using the GGA PBE [44] functional to describe the exchange-correlations interactions. We use the QUICK-

STEP algorithm [45] for the evaluation of the forces as implemented in CP2K [46]. We use the DZVP-MOLOPT-SR-GTH basis set [46, 47] and we employ the Goedecker-Teter-Hutter pseudopotentials [48–50]. The valence orbitals explicitly represented are  $3s^2 3p^6 3d^2 4s^2$  for Ti atoms,  $2s^2 2p^4$  for O atoms and  $2s^2 2p^5$  for F atoms. The plane-wave cutoff is set to 400 Ry. We add dispersive interactions using the DFT-D3 correction [51] and a cutoff radius of 30 Å. We apply isotropic boundary conditions. We make use of the maximally localized Wannier functions (MLWFs) [52] to localize the Kohn and Sham orbitals.

### Classical force-field

Here we present the classical polarizable model employed for the description of the atomic species and we show the parameters that we derived from the DFT simulations. The repulsive and dispersive terms of the interactions are taken into account using the the Born–Mayer–Huggins–Fumi–Tosi–damped (BMHFTD) form of the interaction potential:

$$V_{\text{BMHFTD}} = \sum_{i,j>i} A_{ij} e^{-B_{ij} r_{ij}} - f_6^{ij}(r_{ij}) \frac{C_6^{ij}}{r_{ij}^6} - f_8^{ij}(r_{ij}) \frac{C_8^{ij}}{r_{ij}^8}. \quad (3)$$

The damping functions are Tang-Toennies functions of the form

$$f_n^{ij}(r_{ij}) = 1 - e^{b_D^{ij} r_{ij}} \sum_{k=0}^n \frac{(b_D^{ij} r_{ij})^k}{k!}. \quad (4)$$

The parameters extracted from DFT calculations are reported in Table I.

For the Coulombic part of the interaction potential

$$V_{\text{Coulomb}} = \sum_{i,j>i} \frac{q_i q_j}{r_{ij}}, \quad (5)$$

the formal charges for the ionic species are used,  $-2e$  for O ions,  $-e$  for F ions and  $+4e$  for Ti ions. The many-body electrostatic interactions are described by the induced dipoles  $\boldsymbol{\mu}_i$ , obtained at each MD step minimizing the polarization energy

$$V_{\text{pol}} = \sum_i \frac{1}{2\alpha_i} |\boldsymbol{\mu}_i|^2 + \sum_{i,j>i} [(q^i \mu_\alpha^j g^{ij}(r_{ij}) - q^j \mu_\alpha^i g^{ji}(r_{ij})) T_{ij}^\alpha - \mu_\alpha^i \mu_\beta^j T_{ij}^{\alpha\beta}] \quad (6)$$

where the Einstein summation convention is assumed,  $\alpha_i$  is the atomic polarizability and  $T$  are the multipole

interaction tensors. The damping function  $g_{ij}(r_{ij})$  is of the Tang-Toennies form

$$g_{ij}(r_{ij}) = 1 - c_{ij} e^{-b_{ij} r_{ij}} \sum_{k=0}^4 \frac{(b_{ij} r_{ij})^k}{k!}. \quad (7)$$

The parameters extracted for the polarization part of the potential are reported in Table II.

The classical MD simulations are performed using the software CP2K (step 1, 2 and 3 of the screening procedure, see main text) and the in-house simulation software PIMAIM (step 4 and 5 of the screening procedure, see main text), the short-range interactions are cut off at half the norm of the shortest box vector (or less in *NPT* runs). Isotropic periodic boundary condition are applied. The dipolar Ewald summation method is employed to deal with the electrostatics [53, 54]. The time step for the integration of the equations of motion in the finite temperature runs is 1 fs.



TABLE I. Parameters of the BMHFTD potential extracted from DFT simulations.

Atom Pair	$A_{ij}$ (Ha)	$B_{ij}$ ( $\text{\AA}^{-1}$ )	$C_6^{ij}$ (Ha $\text{\AA}^6$ )	$C_8^{ij}$ (Ha $\text{\AA}^8$ )	$b_D^{ij}$ ( $\text{\AA}^{-1}$ )
O–O	290.4	4.54668	0.48309	2.61949	2.64562
O–F	278.4	4.71487	0.39890	1.55438	3.11805
O–Ti	43.0	2.86431	–	–	–
F–F	282.3	4.61849	0.32938	0.922357	3.59048
F–Ti	28.3	3.13082	–	–	–
Ti–Ti	1.0	9.44863	–	–	–

TABLE II. Parameters of the polarization part of the interaction potential extracted from DFT simulations.

Atom / Atom Pair	$\alpha$ ( $\text{\AA}^3$ )	$b_{ij}$ ( $\text{\AA}^{-1}$ )	$c_{ij}$
O	1.59150		
O–O		4.74888	2.227
O–F		–	–
O–Ti		3.90122	2.13327
F	1.16458		
F–O		–	–
F–F		–	–
F–Ti		4.16887	2.90678
Ti	0.20442		
Ti–O		3.90122	-1.90330
Ti–F		4.16887	-2.66057
Ti–Ti		–	–

- [1] Chen, X.; Mao, S. S., Titanium dioxide nanomaterials: synthesis, properties, modifications, and applications, *Chem. Rev.* **2007**, *107*, 2891–2959.
- [2] Kavan, L.; Grätzel, M.; Gilbert, S. E.; Klemenz, C.; Scheel, H. J., Electrochemical and photoelectrochemical investigation of single-crystal anatase, *J. Am. Chem. Soc.* **1996**, *118*, 6716–6723.
- [3] O'Regan, B.; Grätzel, M., A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO<sub>2</sub> films, *Nature* **1991**, *353*, 737–740.
- [4] Fujishima, A.; Honda, K., Electrochemical photolysis of water at a semiconductor electrode, *Nature* **1972**, *238*, 37–38.
- [5] Hoffmann, M. R.; Martin, S. T.; Choi, W.; Bahnemann, D. W., Environmental applications of semiconductor photocatalysis, *Chem. Rev.* **1995**, *95*, 69–96.
- [6] Kudo, A.; Miseki, Y., Heterogeneous photocatalyst materials for water splitting, *Chem. Soc. Rev.* **2008**, *38*, 253–278.
- [7] Ravelli, D.; Dondi, D.; Fagnoni, M.; Albini, A., Photocatalysis: a multi-faceted concept for green chemistry, *Chem. Soc. Rev.* **2009**, *38*, 1999–2011.
- [8] Kamat, P. V., TiO<sub>2</sub> nanostructures: Recent physical chemistry advances, *J. Phys. Chem. C* **2012**, *116*, 11849–11851.
- [9] Wagemaker, M.; Borghols, W. J. H.; Mulder, F. M., Large impact of particle size on insertion reactions: a case for anatase Li<sub>x</sub>TiO<sub>2</sub>, *J. Am. Chem. Soc.* **2007**, *129*, 4323–4327.
- [10] Yang, M.-C.; Lee, Y.-Y.; Xu, B.; Powers, K.; Meng, Y. S., TiO<sub>2</sub> flakes as anode materials for li-ion-batteries, *J. Power Sources* **2012**, *207*, 166–172.
- [11] Morgan, B. J.; Watson, G. W., Role of lithium ordering in the Li<sub>x</sub>TiO<sub>2</sub> anatase → titanate phase transition, *J. Phys. Chem. Lett.* **2011**, *2*, 1657–1661.
- [12] Yildirim, H.; Greeley, J. P.; Sankaranarayanan, S. K. R. S., Concentration-dependent ordering of lithiated amorphous TiO<sub>2</sub>, *J. Phys. Chem. C* **2013**, *117*, 3834–3845.
- [13] Morgan, B. J.; Madden, P. A., Lithium intercalation into TiO<sub>2</sub>(B) a comparison of LDA, GGA, and GGA+U density functional calculations, *Phys. Rev. B* **2012**, *86*, 035147–035147.
- [14] Morgan, B. J.; Watson, G. W., GGA+U description of lithium intercalation into anatase TiO<sub>2</sub>, *Phys. Rev. B* **2010**, *82*, 144119.
- [15] Tosoni, S.; Lamiel-Garcia, O.; Hevia, D. F.; Illas, F., Theoretical study of atomic fluorine diffusion through bulk TiO<sub>2</sub> polymorphs, *J. Phys. Chem. C* **2013**, *117*, 5855–5860.
- [16] Tosoni, S.; Lamiel-Garcia, O.; Hevia, D. F.; Miguel, J.; Illas, F., Electronic structure of f-doped bulk rutile, anatase, and brookite polymorphs of TiO<sub>2</sub>, *J. Phys. Chem. C* **2012**, *116*, 12738–12746.
- [17] Samadpour, M.; Boix, P. P.; Gimnez, S.; Iraji Zad, A.; Taghavinia, N.; Mora-Ser, I.; Bisquert, J., Fluorine treatment of TiO<sub>2</sub> for enhancing quantum dot sensitized solar cell performance, *J. Phys. Chem. C* **2011**, *115*, 14400–14407.
- [18] Czoska, A. M.; Livraghi, S.; Chiesa, M.; Giamello, E.; Agnoli, S.; Granozzi, G.; Finazzi, E.; Valentin, C. D.; Pacchioni, G., The nature of defects in fluorine-doped TiO<sub>2</sub>, *J. Phys. Chem. C* **2008**, *112*, 8951–8956.
- [19] Dambournet, D.; Chapman, K. W.; Chupas, P. J.; Gerald, R. E.; Penin, N.; Labrugere, C.; Demourgues, A.; Tressaud, A.; Amine, K., Dual lithium insertion and conversion mechanisms in a titanium-based mixed-anion nanocomposite, *J. Am. Chem. Soc.* **2011**, *133*, 13240–13243.
- [20] Wang, Y.; Zhang, H.; Han, Y.; Liu, P.; Yao, X.; Zhao, H., A selective etching phenomenon on {001} faceted anatase titanium dioxide single crystal surfaces by hydrofluoric acid, *Chem. Commun.* **2011**, *47*, 2829–2831.
- [21] Wang, Y.; Zhang, H.; Liu, P.; Sun, T.; Li, Y.; Yang, H.; Yao, X.; Zhao, H., Nature of visible-light responsive fluorinated titanium dioxides, *J. Mater. Chem. A* **2013**, *1*, 12948.
- [22] Yang, H. G.; Sun, C. H.; Qiao, S. Z.; Zou, J.; Liu, G.; Smith, S. C.; Cheng, H. M.; Lu, G. Q., Anatase TiO<sub>2</sub> single crystals with a large percentage of reactive facets, *Nature* **2008**, *453*, 638–641.
- [23] Yang, S.; Halliburton, L. E., Fluorine donors and Ti<sup>3+</sup> ions in TiO<sub>2</sub> crystals, *Phys. Rev. B* **2010**, *81*, 035204–035204.
- [24] Li, W.; Corradini, D.; Body, M.; Legein, C.; Salanne, M.; Chapman, K. W.; Chupas, P. J.; Rollet, A.-L.; Julien, C.; Zhagib, K.; Duttine, M.; Demourgues, A.; Groult, H.; Dambournet, D., Highly fluorinated anatase (TiO<sub>2</sub>) with cationic vacancies **2014**, to be submitted, patent pending (2014).
- [25] Curtarolo, S.; Hart, G. L. W.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O., The high-throughput highway to computational materials design, *Nat. Mater.* **2013**, *12*, 191–201.
- [26] Zakutayev, A.; Zhang, X.; Nagaraja, A.; Yu, L.; Lany, S.; Mason, T. O.; Ginley, D. S.; Zunger, A., Theoretical prediction and experimental realization of new stable inorganic materials using the inverse design approach, *J. Am. Chem. Soc.* **2013**, *135*, 10048–10054.
- [27] Rondinelli, J. M.; Benedek, N. A.; Freedman, D. E.; Kavner, A.; Rodriguez, E. E.; Toberer, E. S.; Martin, L. W., Accelerating functional materials discovery, *Am. Ceram. Soc. Bull.* **2013**, *92*, 14–21.
- [28] Catlow, C. R. A.; Freeman, C. M.; Royle, R. L., Recent studies using static simulation techniques, *Physica B* **1985**, *131*, 1–12.
- [29] Catlow, C. R. A.; James, R., Disorder in TiO<sub>2-x</sub>, *Proc. R. Soc. Lond. A* **1982**, *384*, 157–173.
- [30] Mostoller, M.; Wang, J. C., Ionic potential models in insulators having the rutile structure, *Phys. Rev. B* **1985**, *32*, 6773–6786.
- [31] Matsui, M.; Akaogi, M., Molecular dynamics simulation of the structural and physical properties of the four polymorphs of TiO<sub>2</sub>, *Mol. Simulat.* **1991**, *6*, 239–244.
- [32] Sawatari, H.; Iguchi, E.; Tilley, R., Formation energies of point defects in rutile TiO<sub>2</sub>, *Journal of Physics and Chemistry of Solids* **1982**, *43*, 1147–1155.
- [33] Post, J. E.; Burnham, C. W., Ionic modeling of mineral structures and energies in the electron gas approximation: TiO<sub>2</sub> polymorphs, quartz forsterite, diopside, *Amer. Miner.* **1986**, *71*, 142–150.
- [34] Han, X. J.; Bergqvist, L.; Dederichs, P. H.; Müller-

- Krumbhaar, H.; Christie, J. K.; Scandolo, S.; Tangney, P., Polarizable interatomic force field for  $\text{TiO}_2$  parametrized using density functional theory, *Phys. Rev. B* **2010**, *81*, 134108.
- [35] Salanne, M.; Rotenberg, B.; Jahn, S.; Vuilleumier, R.; Simon, C.; Madden, P. A., Including many-body effects in models for ionic liquids, *Theor. Chem. Acc.* **2012**, *131*, 1–16.
- [36] Tazi, S.; Molina, J. J.; Rotenberg, B.; Turq, P.; Vuilleumier, R.; Salanne, M., A transferable ab initio based force field for aqueous ions, *J. Chem. Phys.* **2012**, *136*, 114507.
- [37] Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q., Large-scale screening of hypothetical metalorganic frameworks, *Nat. Chem.* **2012**, *4*, 83–89.
- [38] Horn, M.; Schwerdtfeger, C. F.; Meagher, E. P., Refinement of the structure of anatase at several temperatures, *Z. Kristallogr.* **1972**, *136*, 273–281.
- [39] Nosé, S., A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.* **1984**, *52*, 255–268.
- [40] Hoover, W. G., Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A* **1985**, *31*, 1695–1697.
- [41] Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms, *J. Chem. Phys.* **1994**, *101*, 4177–4189.
- [42] Hemmati, M.; Wilson, M.; Madden, P. A., Structure of liquid  $\text{Al}_2\text{O}_3$  from a computer simulation model, *J. Phys. Chem. B* **1999**, *103*, 4023–4028.
- [43] Wilson, A. J. C., editor, *International tables for X-ray crystallography*, Kluwer Academic Publisher, Dordrecht, **1992**.
- [44] Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized gradient approximation made simple, *Phys. Rev. Lett.* **1996**, *77*, 3865.
- [45] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J., QUICKSTEP: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach, *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- [46] CP2K developers group. See <http://www.cp2k.org>.
- [47] VandeVondele, J.; Hutter, J., Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, *J. Chem. Phys.* **2007**, *127*, 114105–114109.
- [48] Goedecker, S.; Teter, M.; Hutter, J., Separable dual-space gaussian pseudopotentials, *Phys. Rev. B* **1996**, *54*, 1703–1710.
- [49] Hartwigsen, C.; Goedecker, S.; Hutter, J., Relativistic separable dual-space gaussian pseudopotentials from H to Rn, *Phys. Rev. B* **1998**, *58*, 3641–3662.
- [50] Krack, M., Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals, *Theor. Chem. Acc.* **2005**, *114*, 145–152.
- [51] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.* **2010**, *132*, 154104.
- [52] Martin, R., *Electronic structure : basic theory and practical methods*, Cambridge University Press, Cambridge, UK, **2004**.
- [53] Aguado, A.; Madden, P. A., Ewald summation of electrostatic multipole interactions up to the quadrupolar level, *J. Chem. Phys.* **2003**, *119*, 7471–7483.
- [54] Laino, T.; Hutter, J., Notes on “Ewald summation of electrostatic multipole interactions up to quadrupolar level” [J. Chem. Phys. 119, 7471 (2003)], *J. Chem. Phys.* **2008**, *129*, 074102.